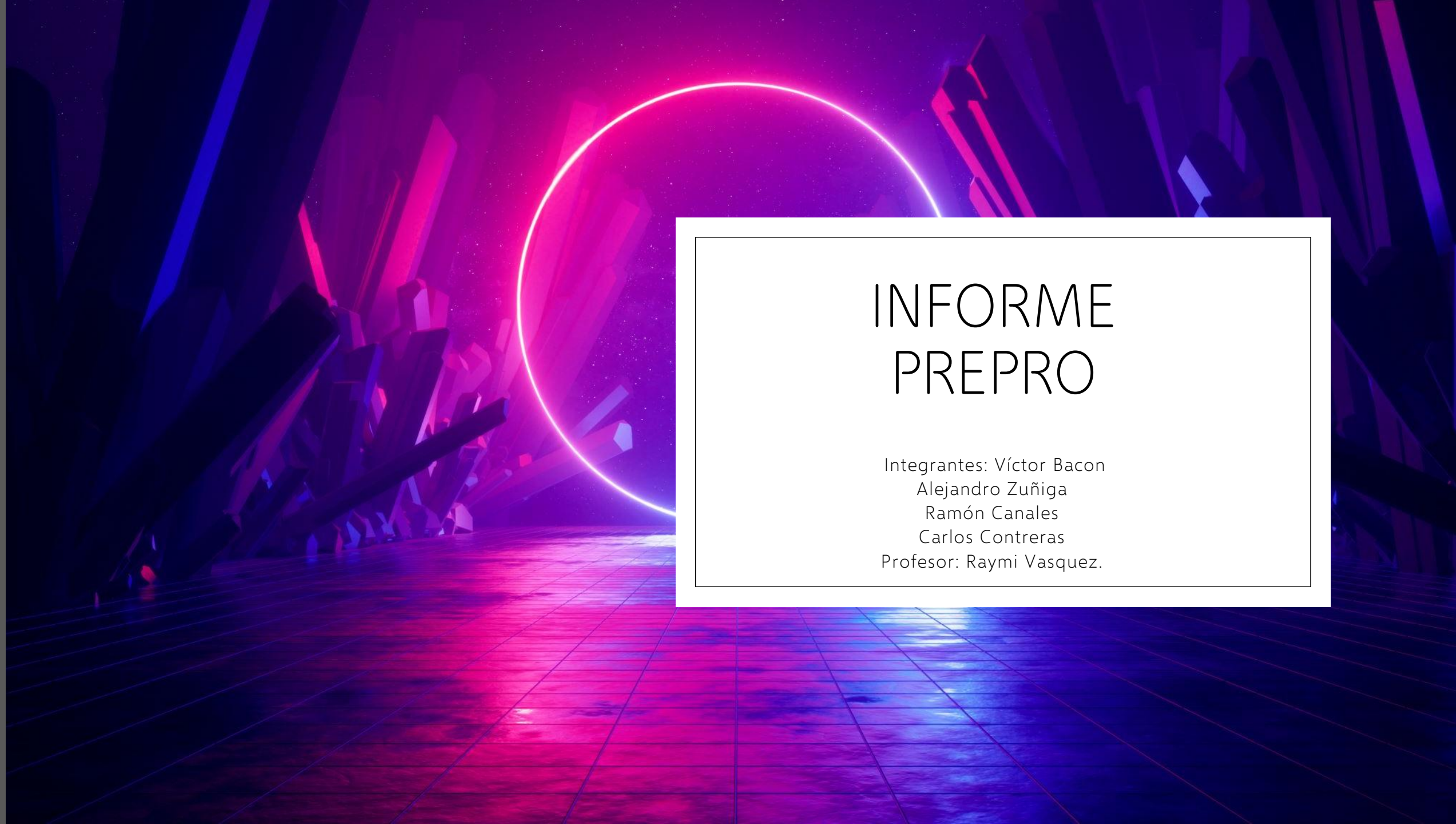
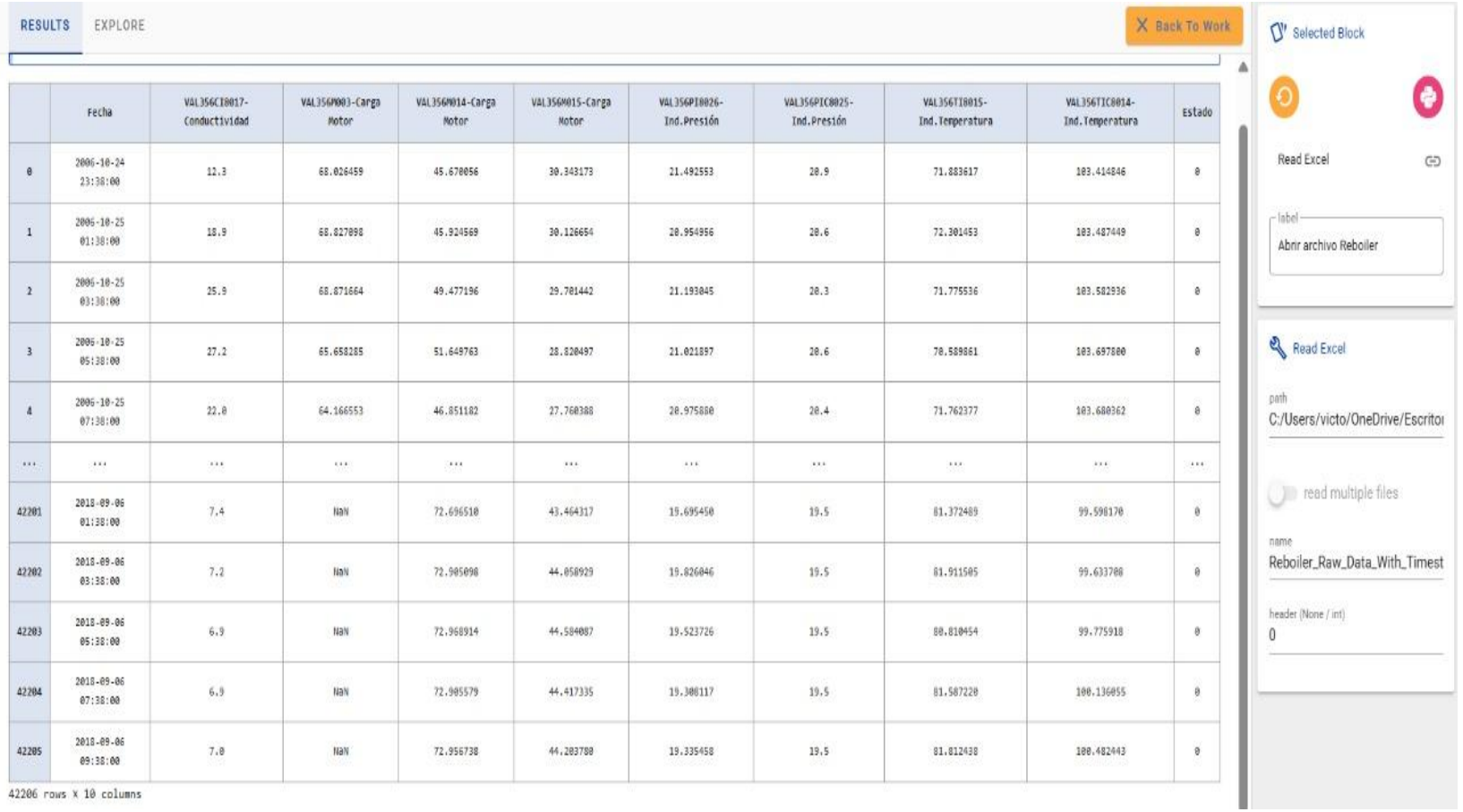


INFORME PREPRO

Integrantes: Víctor Bacon
Alejandro Zuñiga
Ramón Canales
Carlos Contreras
Profesor: Raymi Vasquez.

PASO 1
READ EXCEL

- Crear un Bloque lector de Excel donde se debe cargar la base de datos dadas por el profesor.
- El header debe ser 0 para que el encabezado se organice según la fecha.

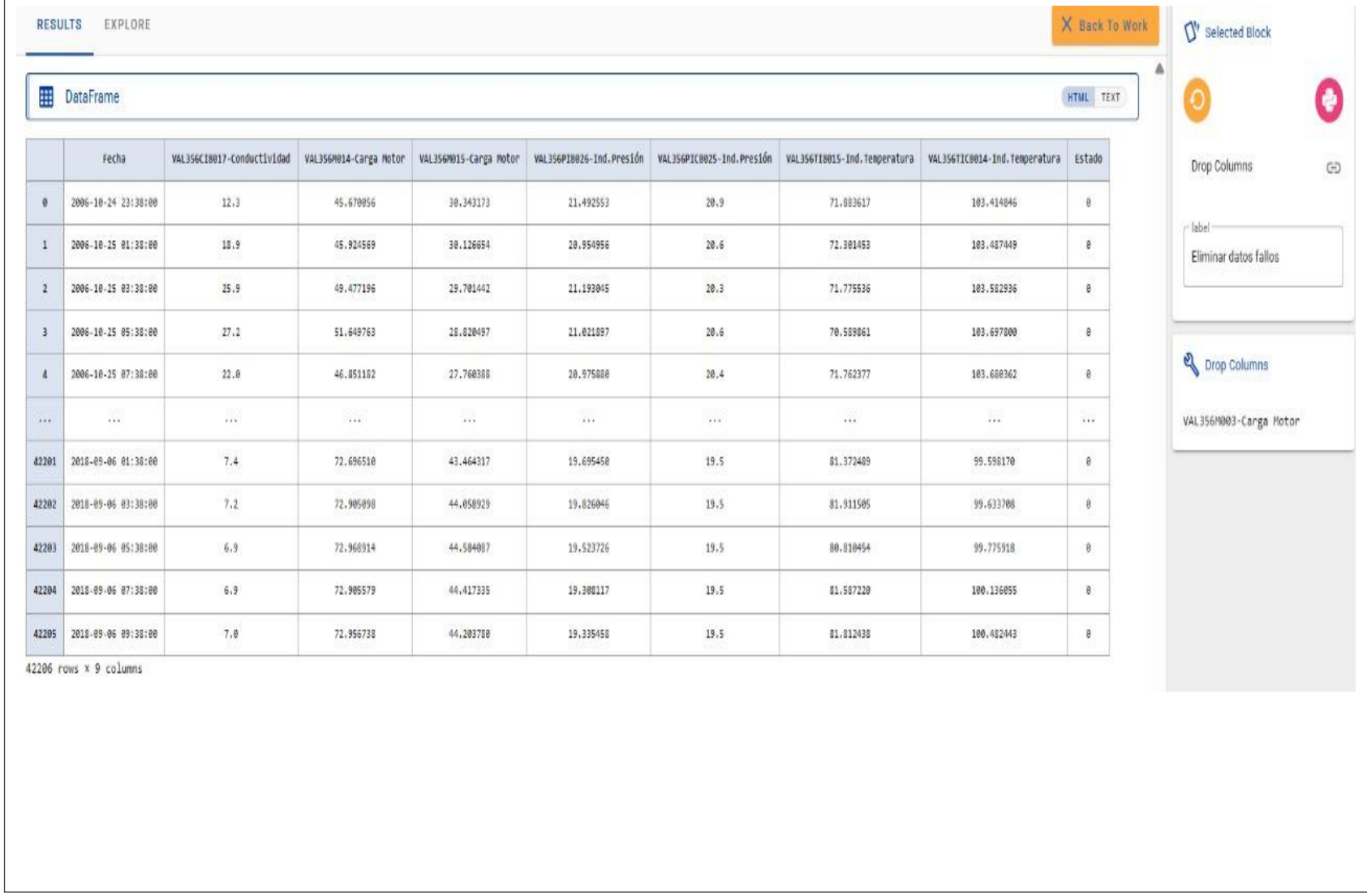


	Fecha	VAL356CI8017-Conductividad	VAL356M003-Carga Motor	VAL356M014-Carga Motor	VAL356M015-Carga Motor	VAL356PI8026-Ind.Presión	VAL356PIC8025-Ind.Presión	VAL356TI8015-Ind.Temperatura	VAL356TIC8014-Ind.Temperatura	Estado
0	2006-10-24 23:38:00	12.3	68.026459	45.670056	30.343173	21.492553	20.9	71.883617	103.414846	0
1	2006-10-25 01:38:00	18.9	68.827898	45.924569	30.126654	20.954956	20.6	72.301453	103.487449	0
2	2006-10-25 03:38:00	25.9	68.871664	49.477196	29.701442	21.193845	20.3	71.775536	103.582936	0
3	2006-10-25 05:38:00	27.2	65.658285	51.648763	28.820497	21.021897	20.6	70.589861	103.697800	0
4	2006-10-25 07:38:00	22.0	64.166553	46.851182	27.760388	20.975880	20.4	71.762377	103.680362	0
...	...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	NaN	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	NaN	72.905098	44.050929	19.826046	19.5	81.911505	99.633708	0
42203	2018-09-06 05:38:00	6.9	NaN	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	NaN	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	NaN	72.956738	44.283788	19.335458	19.5	81.812438	100.482443	0

42206 rows x 10 columns

PASO 2
DROP COLUMNS

- Eliminar datos de fallas, en este paso se filtran los datos que no son necesarios para el análisis.

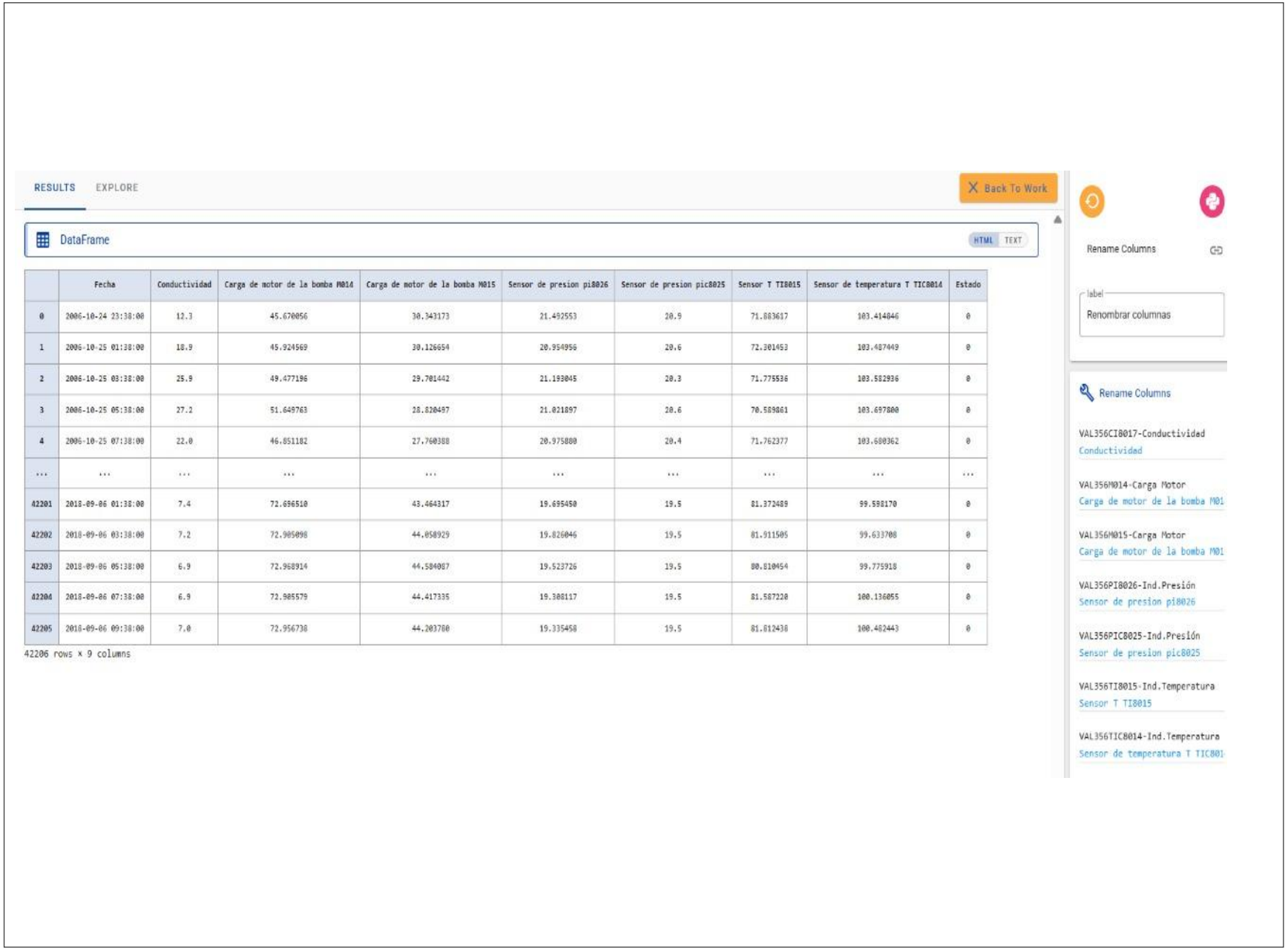


	Fecha	VAL356CI8017-Conductividad	VAL356M014-Carga Motor	VAL356M015-Carga Motor	VAL356PI8026-Ind.Presión	VAL356PIC8025-Ind.Presión	VAL356TI8015-Ind.Temperatura	VAL356TIC8014-Ind.Temperatura	Estado
0	2006-10-24 23:38:00	12.3	45.670056	30.343173	21.492553	20.9	71.883617	103.414846	0
1	2006-10-25 01:38:00	18.9	45.924569	30.126654	20.954956	20.6	72.301453	103.487449	0
2	2006-10-25 03:38:00	25.9	49.477196	29.701442	21.193045	20.3	71.775536	103.582936	0
3	2006-10-25 05:38:00	27.2	51.649763	28.820497	21.021897	20.6	70.589861	103.697800	0
4	2006-10-25 07:38:00	22.0	46.851182	27.760388	20.975888	20.4	71.762377	103.688362	0
...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695458	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	72.905098	44.058929	19.826046	19.5	81.911505	99.633708	0
42203	2018-09-06 05:38:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0

42206 rows x 9 columns

PASO 3
RENAME COLUMNS

- Se renombran columnas para tener una mejor visual de la información y así generar mayor entendimiento.

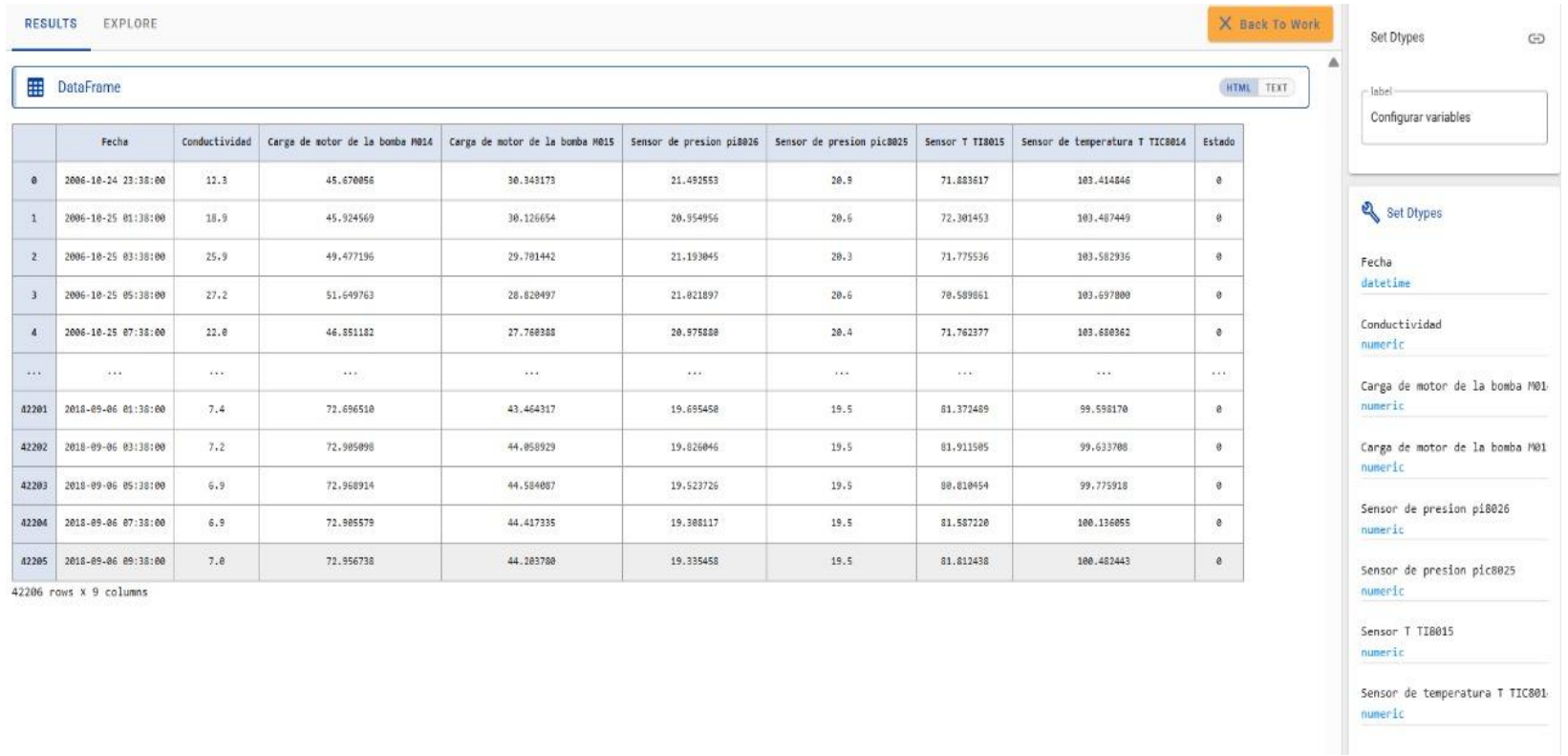


	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
0	2006-10-24 23:38:00	12.3	45.670056	30.343173	21.492553	20.9	71.883617	103.414046	0
1	2006-10-25 01:38:00	18.9	45.924569	30.126654	20.954956	20.6	72.301453	103.487449	0
2	2006-10-25 03:38:00	25.9	49.477196	29.701442	21.193045	20.3	71.775536	103.582936	0
3	2006-10-25 05:38:00	27.2	51.648763	28.820497	21.021897	20.6	70.589861	103.697800	0
4	2006-10-25 07:38:00	22.0	46.851182	27.760388	20.975800	20.4	71.762377	103.680362	0
...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	72.905098	44.058929	19.826046	19.5	81.911505	99.633708	0
42203	2018-09-06 05:38:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	72.985579	44.417335	19.388117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0

42206 rows x 9 columns

PASO 4
SET DTYPES

- Configurar variables y se separan los datos numéricos con los datos de fecha y hora, variables que son requeridas para el análisis.

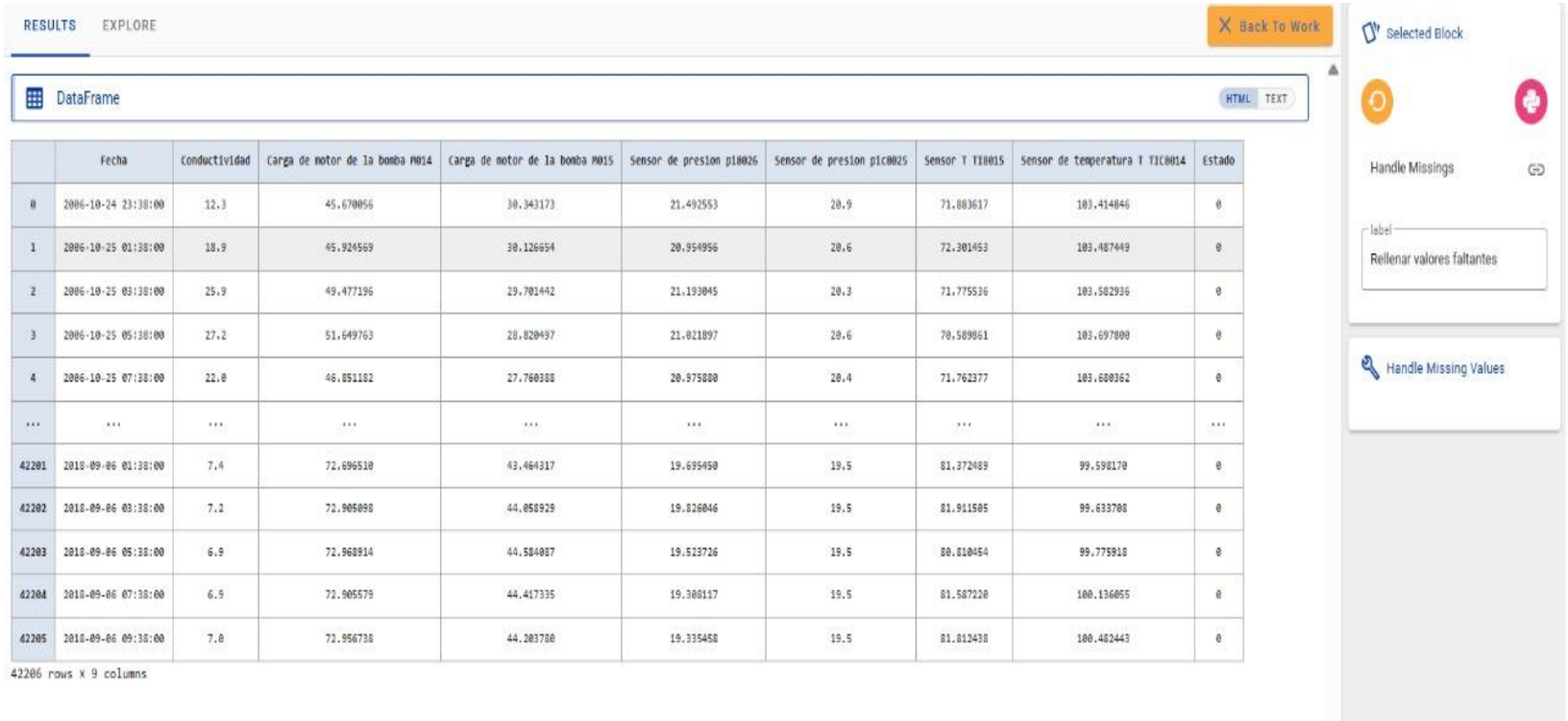


	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
0	2006-10-24 23:38:00	12.3	45.670856	30.343173	21.492553	20.9	71.883617	103.414846	0
1	2006-10-25 01:38:00	18.9	45.924569	30.126654	20.954956	20.6	72.301453	103.487449	0
2	2006-10-25 03:38:00	25.9	49.477196	29.701442	21.193045	20.3	71.775536	103.582936	0
3	2006-10-25 05:38:00	27.2	51.649763	28.820497	21.021897	20.6	70.589861	103.697800	0
4	2006-10-25 07:38:00	22.0	46.851182	27.760388	20.975888	20.4	71.762377	103.680362	0
...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695458	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	72.905898	44.058929	19.826046	19.5	81.911585	99.633708	0
42203	2018-09-06 05:38:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	72.985579	44.417335	19.388117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	72.956738	44.283700	19.335458	19.5	81.812438	100.482443	0

42206 rows x 9 columns

PASO 5
HEADLE MISING

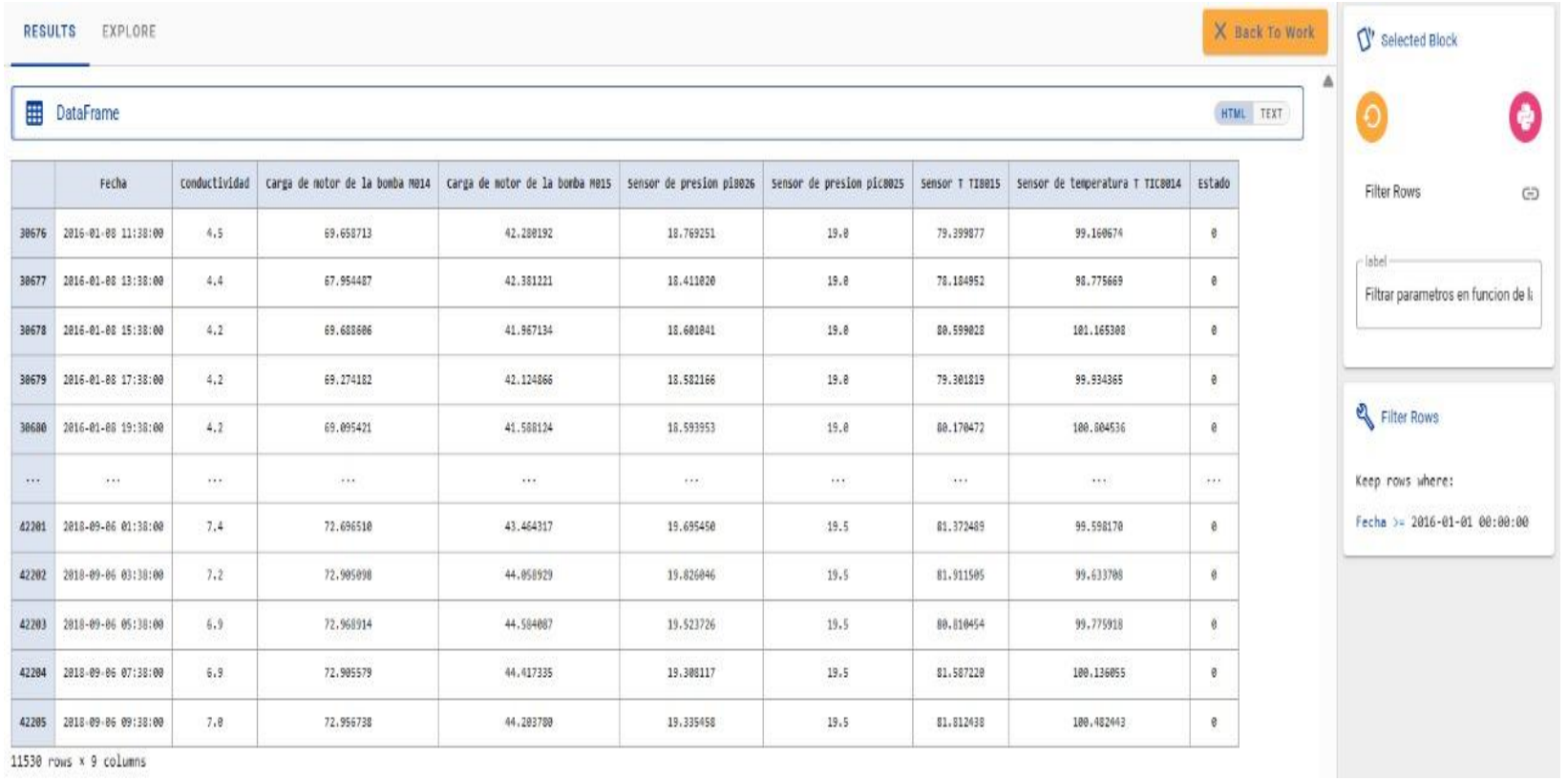
- Se rellenan valores faltantes nulos no registrados en el análisis.



	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion p18026	Sensor de presion p1c8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
0	2006-10-24 23:38:00	12.3	45.670056	30.343173	21.492553	20.9	71.883617	103.414046	0
1	2006-10-25 01:38:00	18.9	45.924569	30.126654	20.954956	20.6	72.301453	103.487449	0
2	2006-10-25 03:38:00	25.9	49.477196	29.701442	21.193045	20.3	71.775536	103.582936	0
3	2006-10-25 05:38:00	27.2	51.649763	28.820497	21.021897	20.6	70.589061	103.697800	0
4	2006-10-25 07:38:00	22.0	46.851182	27.760388	20.975880	20.4	71.762377	103.680362	0
...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	72.905098	44.058929	19.826046	19.5	81.911505	99.633708	0
42203	2018-09-06 05:38:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	72.905579	44.417335	19.388117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	72.956738	44.283780	19.335458	19.5	81.812438	100.482443	0

PASO 6
FILTER ROWES

- Se filtran los datos según rubrica especificada por el profesor, el cual dejamos los datos desde el 2016 en adelante.

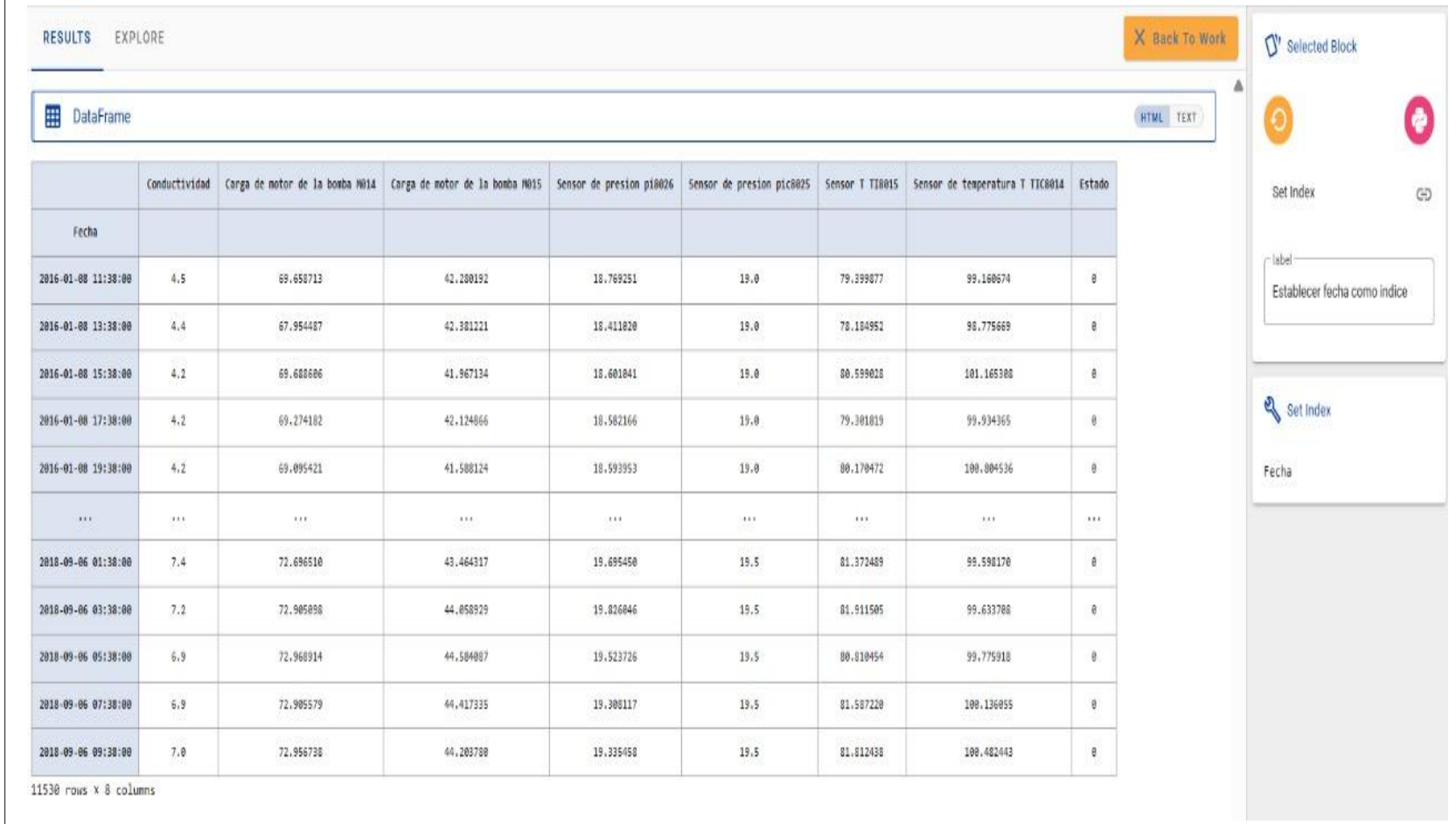


	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
30676	2016-01-08 11:38:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0
30677	2016-01-08 13:38:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0
30678	2016-01-08 15:38:00	4.2	69.688686	41.967134	18.601041	19.0	80.599028	101.165308	0
30679	2016-01-08 17:38:00	4.2	69.274182	42.124866	18.582166	19.0	79.301819	99.934365	0
30680	2016-01-08 19:38:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.804536	0
...	...	...	...	...	...	...	...	...	...
42201	2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
42202	2018-09-06 03:38:00	7.2	72.905090	44.058929	19.826046	19.5	81.911505	99.633708	0
42203	2018-09-06 05:38:00	6.9	72.960914	44.584087	19.523726	19.5	80.810454	99.775918	0
42204	2018-09-06 07:38:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0
42205	2018-09-06 09:38:00	7.0	72.956738	44.283700	19.335458	19.5	81.812438	100.482443	0

11530 rows x 9 columns

PASO 7
SET INDEX

- Establecer fecha como índice para cada fila de los dato a evaluar.

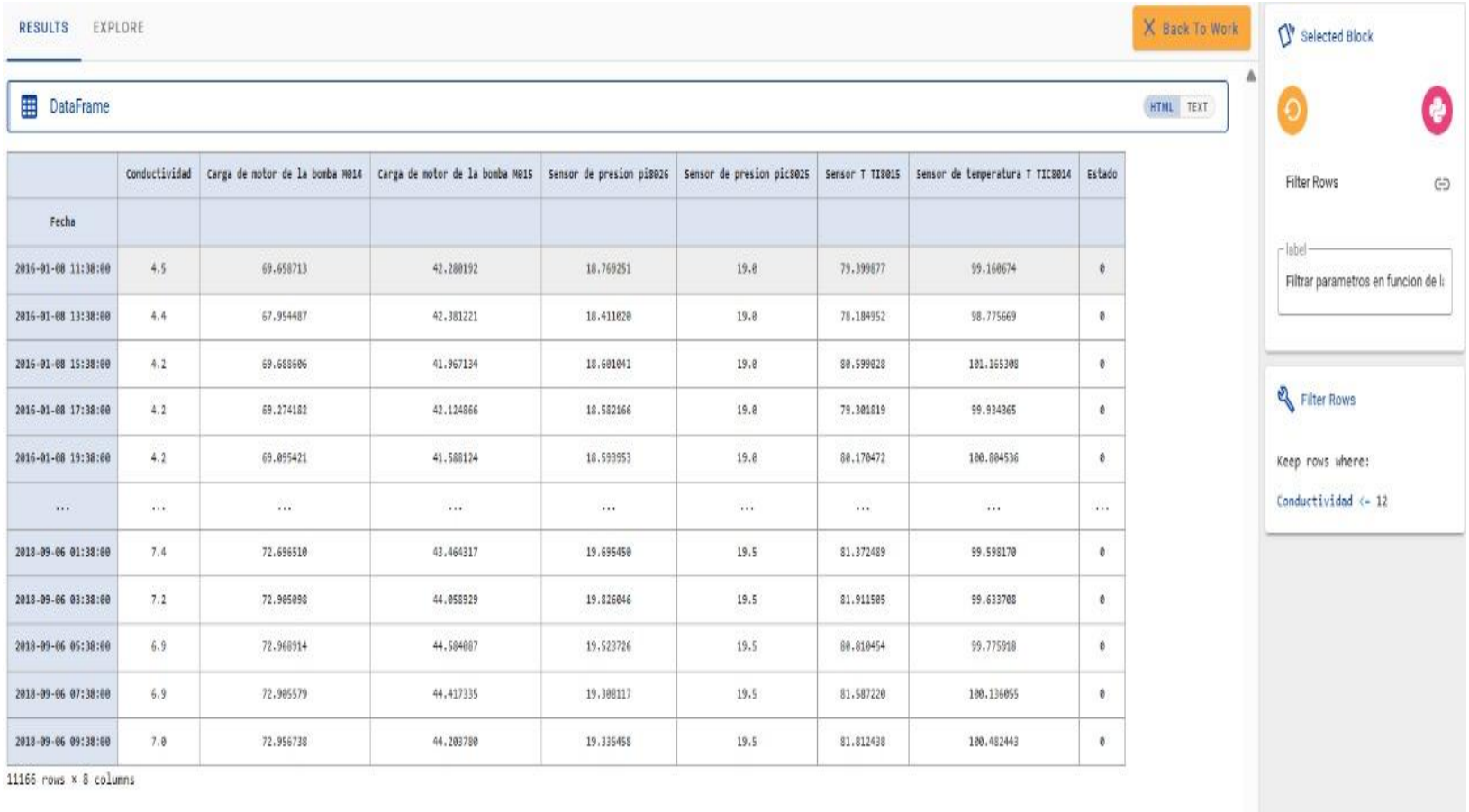


	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
Fecha								
2016-01-08 11:38:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0
2016-01-08 13:38:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0
2016-01-08 15:38:00	4.2	69.688686	41.967134	18.601041	19.0	80.599028	101.165308	0
2016-01-08 17:38:00	4.2	69.274102	42.124866	18.582166	19.0	79.301819	99.934365	0
2016-01-08 19:38:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.804536	0
...	...	...	...	...	...	...	...	...
2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
2018-09-06 03:38:00	7.2	72.905898	44.058929	19.826046	19.5	81.911505	99.633708	0
2018-09-06 05:38:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0
2018-09-06 07:38:00	6.9	72.985579	44.417335	19.308117	19.5	81.587220	100.136055	0
2018-09-06 09:38:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0

11530 rows x 8 columns

PASO 8
FILTER ROWS

- Según lo pedido por el profesor en este análisis se incluye la eliminación de outlier del sensor de conductividad, el cual deben ser eliminados en el data cuando la conductividad es superior a 12 microsiemens/cm.

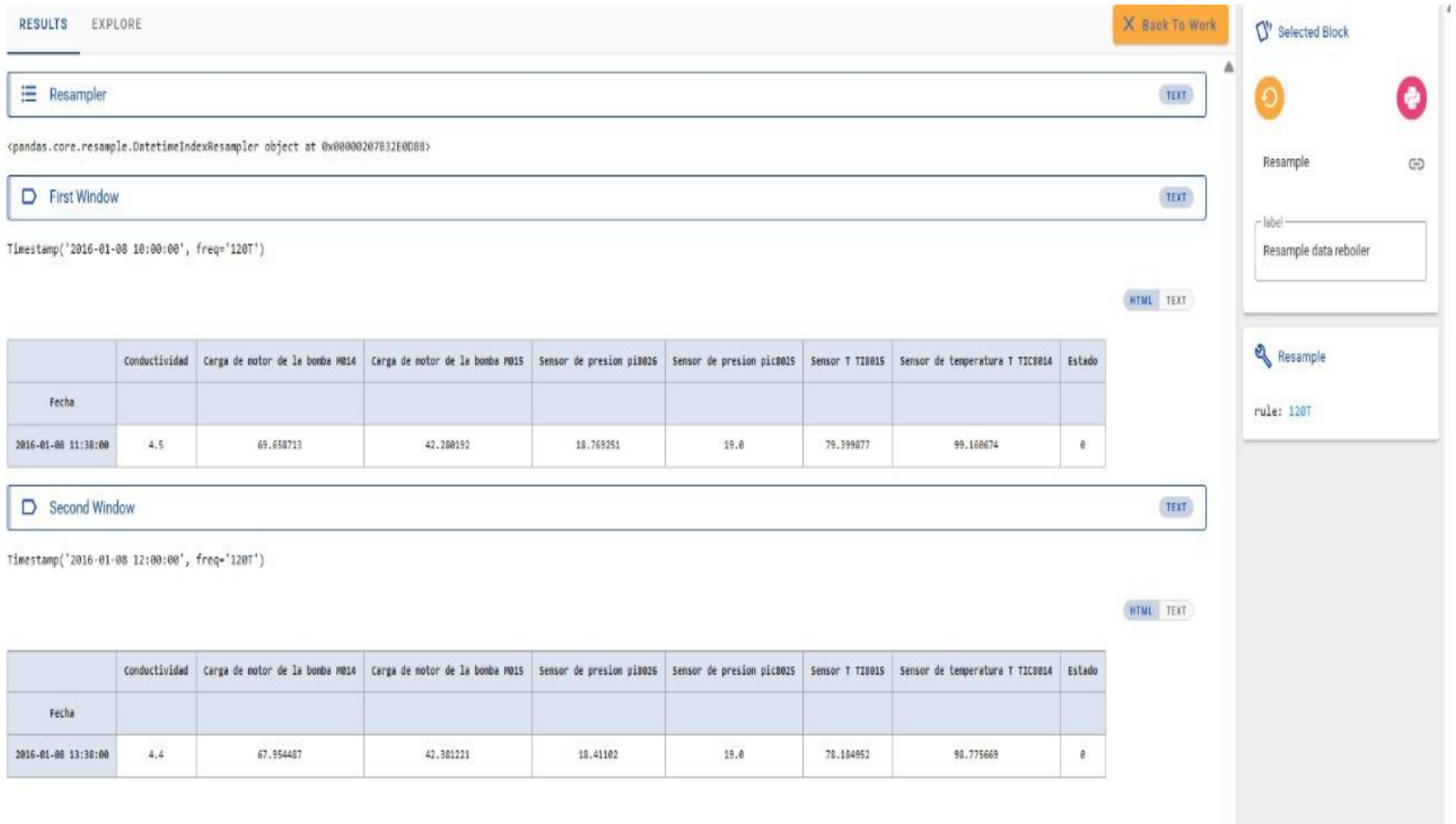


Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
2016-01-08 11:38:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0
2016-01-08 13:38:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0
2016-01-08 15:38:00	4.2	69.688606	41.967134	18.601041	19.0	80.599028	101.165308	0
2016-01-08 17:38:00	4.2	69.274182	42.124866	18.582166	19.0	79.301819	99.934365	0
2016-01-08 19:38:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.884536	0
...	...	...	...	...	...	...	...	...
2018-09-06 01:38:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0
2018-09-06 03:38:00	7.2	72.905898	44.053929	19.826046	19.5	81.911585	99.633708	0
2018-09-06 05:38:00	6.9	72.960914	44.584087	19.523726	19.5	80.810454	99.775918	0
2018-09-06 07:38:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0
2018-09-06 09:38:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0

11166 rows x 8 columns

PASO 9
RESAMPLE

- Se Configura una celda para remuestreo disponible establecida en el lapso de 120 minutos (2horas).

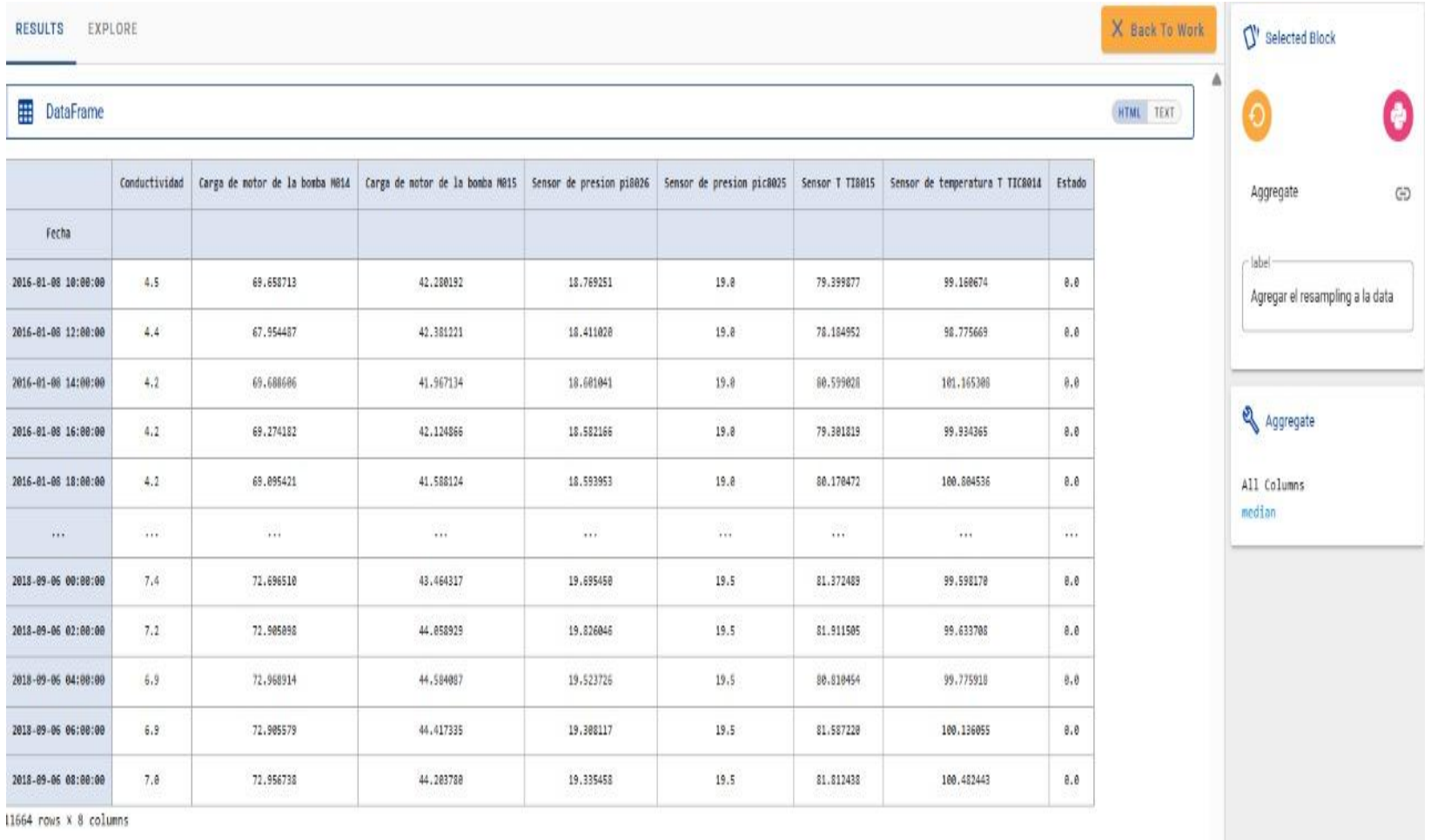


	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
Fecha								
2016-01-08 11:38:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.168674	0

	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
Fecha								
2016-01-08 13:38:00	4.4	67.954487	42.381221	18.41102	19.0	78.184952	98.775669	0

PASO 10
AGREGATE

- Luego de realizar el remuestreo, se agrega un bloque llamado agregate ya que siempre tiene que ir después del remuestreo, el cual tiene diferentes funciones estadísticas que realizan un resumen sobre columnas de datos.

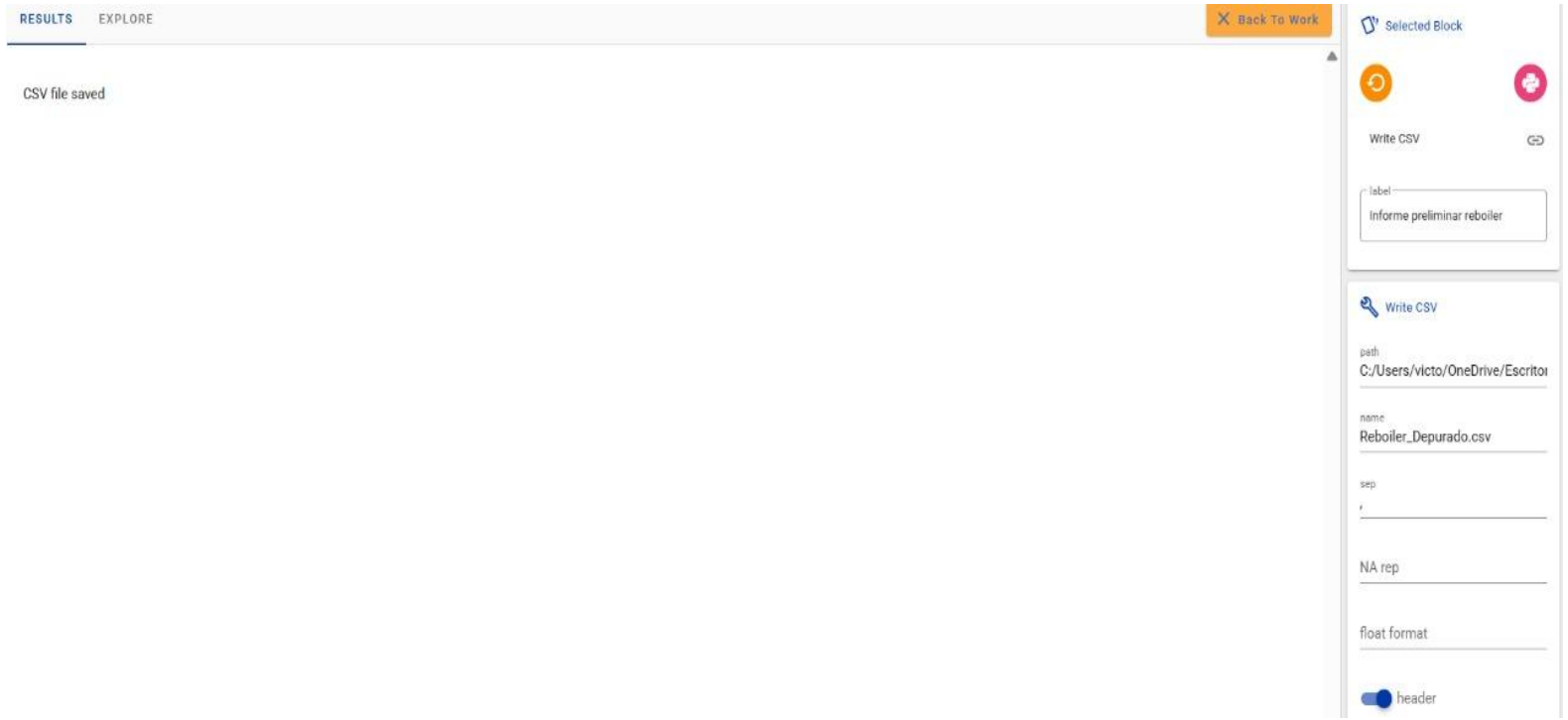


	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
Fecha								
2016-01-08 10:00:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0.0
2016-01-08 12:00:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0.0
2016-01-08 14:00:00	4.2	69.688606	41.967134	18.601041	19.0	80.599020	101.165300	0.0
2016-01-08 16:00:00	4.2	69.274182	42.124866	18.582166	19.0	79.381819	99.934365	0.0
2016-01-08 18:00:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.804536	0.0
...	...	...	...	...	...	...	...	...
2018-09-06 00:00:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598178	0.0
2018-09-06 02:00:00	7.2	72.905898	44.058929	19.826046	19.5	81.911505	99.633708	0.0
2018-09-06 04:00:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775910	0.0
2018-09-06 06:00:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0.0
2018-09-06 08:00:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0.0

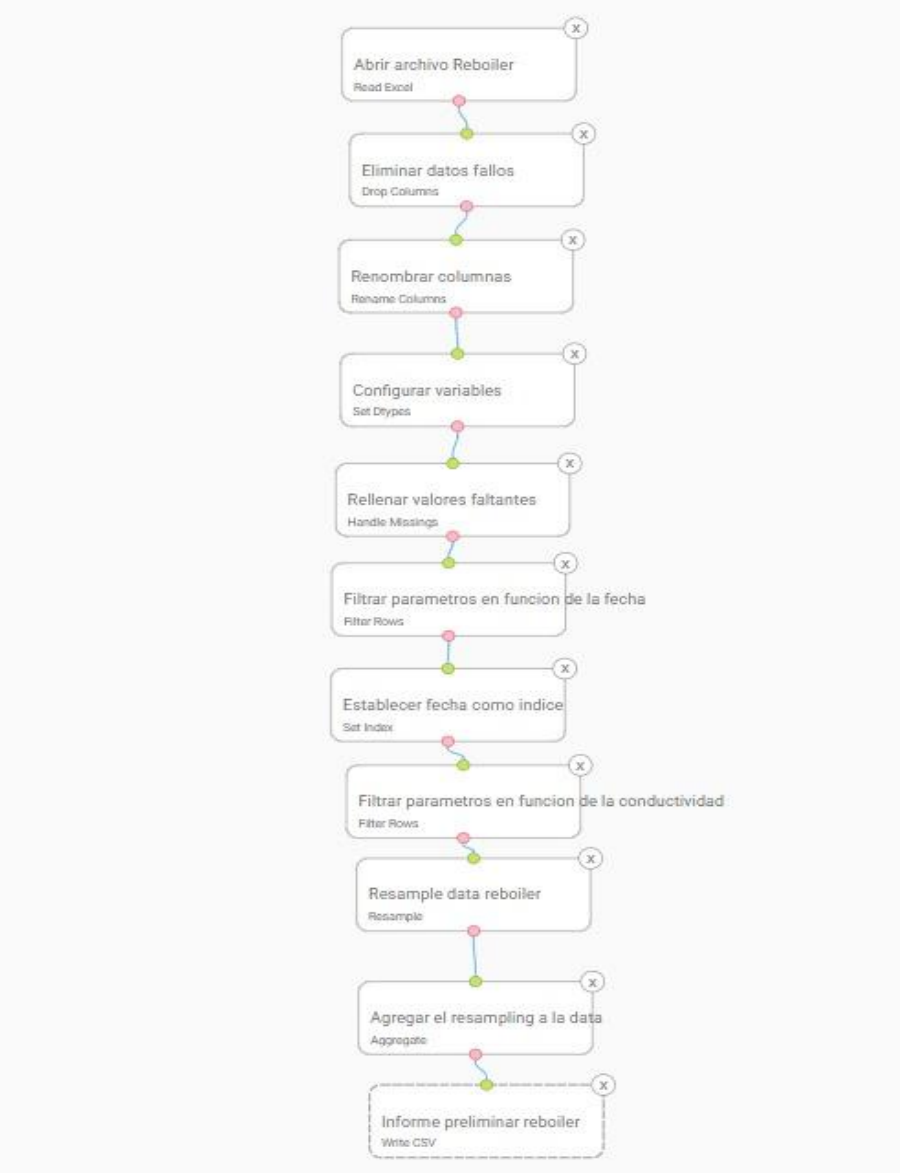
11664 rows x 8 columns

PASO 11
READ CSV

- Se incorpora el bloque que nos permite descargar el archivo CSV.



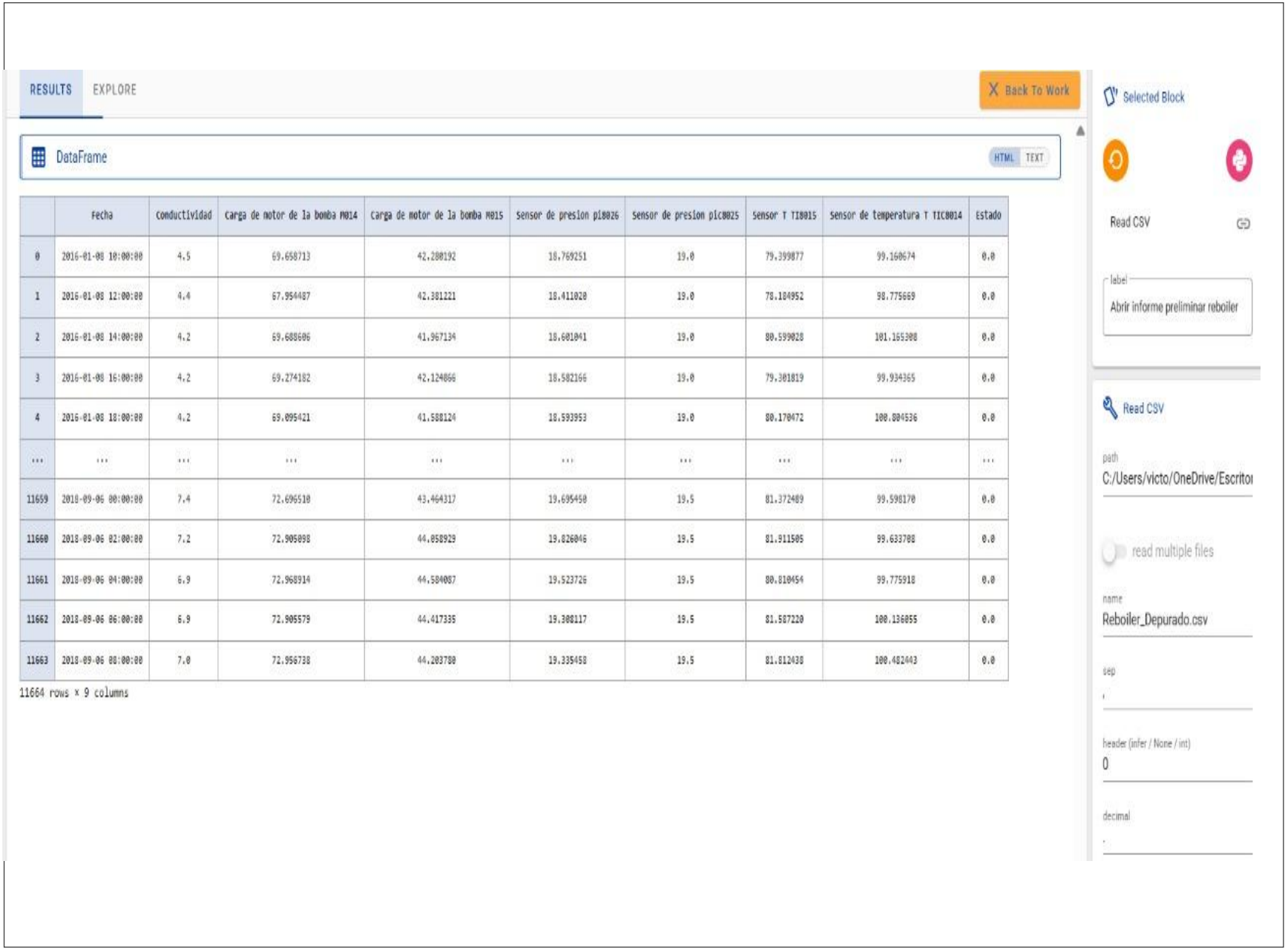
- De esta forma Nuestro primer encadenado queda con la siguiente distribución.



PASO 12

READ CSV

Se realiza otro bloque el cual nos permite descargar Nuestro reboiler depurado.

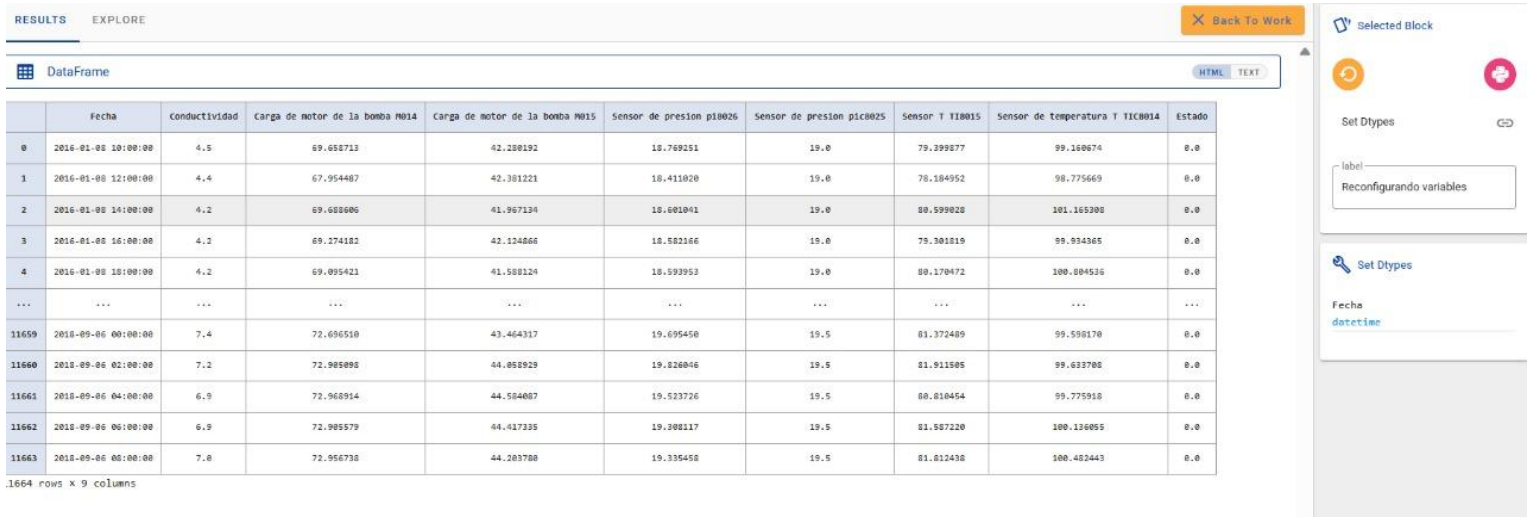


	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
0	2016-01-08 10:00:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0.0
1	2016-01-08 12:00:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0.0
2	2016-01-08 14:00:00	4.2	69.688606	41.967134	18.601041	19.0	80.599028	101.165308	0.0
3	2016-01-08 16:00:00	4.2	69.274182	42.124066	18.582166	19.0	79.301819	99.934365	0.0
4	2016-01-08 18:00:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.804536	0.0
...	...	...	...	...	...	...	...	...	...
11659	2018-09-06 00:00:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0.0
11660	2018-09-06 02:00:00	7.2	72.905898	44.858929	19.826046	19.5	81.911505	99.633708	0.0
11661	2018-09-06 04:00:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0.0
11662	2018-09-06 06:00:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0.0
11663	2018-09-06 08:00:00	7.0	72.956738	44.203780	19.335458	19.5	81.812438	100.482443	0.0

11664 rows × 9 columns

PASO 13
SET DTYPES

- Se reconfiguran variables como fecha y hora.

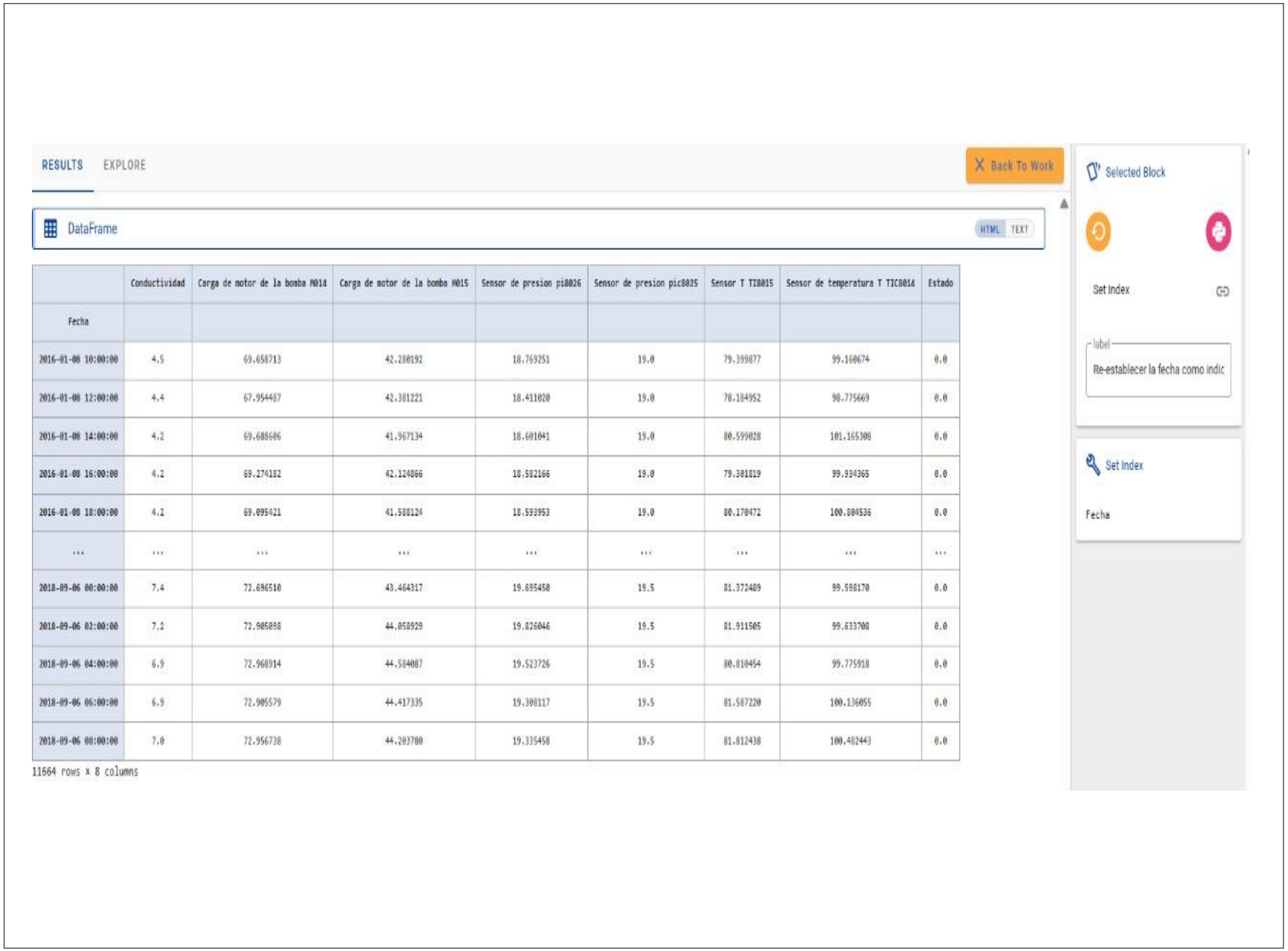


	Fecha	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion p10026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
0	2016-01-08 10:00:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0.0
1	2016-01-08 12:00:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0.0
2	2016-01-08 14:00:00	4.2	69.688606	41.967134	18.601041	19.0	80.599028	101.165308	0.0
3	2016-01-08 16:00:00	4.2	69.274181	42.124866	18.582166	19.0	79.301819	99.934365	0.0
4	2016-01-08 18:00:00	4.2	69.095421	41.588124	18.593953	19.0	80.170472	100.804536	0.0
...	...	...	...	...	...	...	...	...	...
11659	2018-09-06 00:00:00	7.4	72.696510	43.464317	19.695450	19.5	81.372489	99.598170	0.0
11660	2018-09-06 02:00:00	7.2	72.985098	44.058929	19.826046	19.5	81.911505	99.633708	0.0
11661	2018-09-06 04:00:00	6.9	72.968914	44.584087	19.523726	19.5	80.810454	99.775918	0.0
11662	2018-09-06 06:00:00	6.9	72.985579	44.417335	19.308117	19.5	81.587220	100.136055	0.0
11663	2018-09-06 08:00:00	7.0	72.956738	44.283788	19.335458	19.5	81.812438	100.482443	0.0

1664 rows x 9 columns

PASO 14
SET INDEX

- Se reestablece fecha como índice.

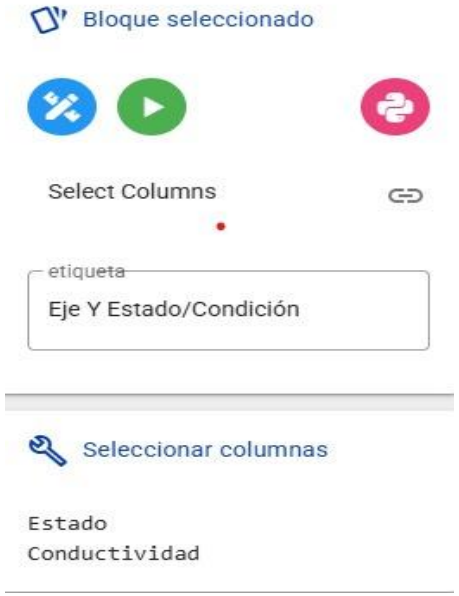


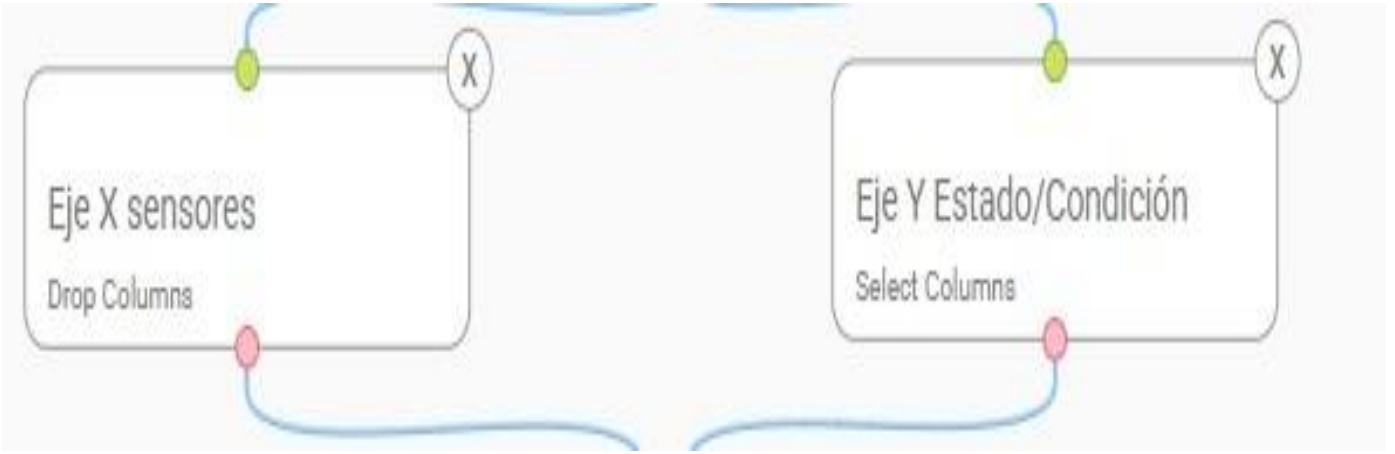
	Conductividad	Carga de motor de la bomba M014	Carga de motor de la bomba M015	Sensor de presion pi8026	Sensor de presion pic8025	Sensor T TI8015	Sensor de temperatura T TIC8014	Estado
Fecha								
2016-01-08 10:00:00	4.5	69.658713	42.280192	18.769251	19.0	79.399877	99.160674	0.0
2016-01-08 12:00:00	4.4	67.954487	42.381221	18.411020	19.0	78.184952	98.775669	0.0
2016-01-08 14:00:00	4.2	69.688686	41.967134	18.601041	19.0	80.599028	101.165308	0.0
2016-01-08 16:00:00	4.2	69.274182	42.124866	18.582166	19.0	79.381819	99.934365	0.0
2016-01-08 18:00:00	4.2	69.095421	41.588124	18.593953	19.0	80.178472	100.804536	0.0
...	...	...	...	...	...	...	...	...
2018-09-06 00:00:00	7.4	72.696510	43.464317	19.695458	19.5	81.372489	99.598170	0.0
2018-09-06 02:00:00	7.2	72.905898	44.058929	19.826046	19.5	81.911505	99.633708	0.0
2018-09-06 04:00:00	6.9	72.960914	44.584087	19.523726	19.5	80.810454	99.775918	0.0
2018-09-06 06:00:00	6.9	72.905579	44.417335	19.308117	19.5	81.587220	100.136055	0.0
2018-09-06 08:00:00	7.0	72.956738	44.283780	19.335458	19.5	81.812438	100.482443	0.0

11664 rows x 8 columns

PASO 15 DROP COLUMNS SELECT COLUMNS

Se instalan dos bloques en paralelo con el objetivo que en un bloque se coloquen los sensores a medir y en el segundo bloque el estado/condición.





PASO 15
NORMAL/ANOMAL SPLINT

Se realizan ventanas de 48 hrs para poder tener de forma horizontal lecturas en tiempo de 24 hrs.

RESULTS EXPLORE

shapes

```
{'x_train': '6930 x (24, 7)',
 'x_test': '4618 x (24, 7)',
 'y_test': '4618 x ()'}
```

x train

```
array([[[0.075     , 0.        , 0.        , ..., 0.97244546,
         0.04825028, 0.04985567],
        [0.06666667, 0.        , 0.        , ..., 0.97359357,
         0.04376356, 0.04759148],
        [0.03333333, 0.        , 0.        , ..., 0.97474168,
         0.04005303, 0.04532452],
        ...,
        [0.425     , 0.34688851, 0.        , ..., 0.09184845,
         0.54833319, 0.60526058],
        [0.425     , 0.4392253 , 0.        , ..., 0.09184845,
         0.55303714, 0.60635885],
        [0.425     , 0.88274552, 0.31740714, ..., 0.09184845,
         0.56139741, 0.61327906]],

       [[0.34166667, 0.9168376 , 0.81393625, ..., 0.04707233,
         0.61849149, 0.59986594],
        [0.34166667, 0.91204282, 0.79779197, ..., 0.04707233,
         0.594917  , 0.59271708],
        [0.33333333, 0.91711313, 0.77106577, ..., 0.04707233,
         0.58832519, 0.59853372],
        ...,
        [0.35      , 0.92765454, 0.77573517, ..., 0.05281286,
         0.62741318, 0.60345964],
        [0.35      , 0.92409917, 0.76873375, ..., 0.05281286,
         0.62935807, 0.60665666],
```

PASO 16
WRITE NORMAL/ANORMAL NPZ

Finalmente luego de realizar los filtros requeridos podemos descargar nuestros archivos reboiler NPZ.